

The DNA Binding Domain of the Human c-Abl Tyrosine Kinase Preferentially Binds to DNA Sequences Containing an AAC Motif and to Distorted DNA Structures[†]

Marie-Hélène David-Cordonnier,[‡] Malika Hamdane,^{‡,§} Christian Bailly,^{‡,||} and Jean-Claude D'Halluin^{*,‡}

INSERM U 124 Onco-hématologie Moléculaire, Institut de Recherches sur le Cancer de Lille, Place de Verdun, 59045 Lille, France, and Centre Oscar Lambret, Place de Verdun, 59045 Lille, France

Received December 10, 1997; Revised Manuscript Received February 9, 1998

ABSTRACT: The c-Abl tyrosine kinase protein is implicated in the signaling pathway as well as in transcription, DNA repair, apoptosis, and several other vital biological processes essential for cell proliferation or differentiation. The interaction of c-Abl with DNA is important for some of these functions, but the exact nature of this interaction is still a matter of controversy. The present study addresses the DNA-binding properties of the human c-Abl protein. Using CASTing experiments, the consensus binding site 5'-A^A/cAACAA^A/c was determined. The central highly conserved AAC triplet appears to constitute the crucial core element in the binding sequences of the c-Abl protein. The c-Abl DNA-binding domain recognizes specific sequences and interacts with deformed DNA structures such as four-way junctions and bubble DNA containing a large single-stranded loop, as determined by electromobility shift assay, melting temperature studies, and binding to specific oligonucleotides covalently linked to beads. Additional competition experiments suggest that the interaction mainly involves contacts within the minor groove of the double helix. The DNA-binding properties of c-Abl are reminiscent of those of high-mobility group (HMG)-like proteins such as LEF-1 and SRY. However, the circular permutation and ring closure assays and DNA unwinding experiments reveal that, unlike HMGs, c-Abl does not bend its target sequence. In addition, it is shown that the protein potentiates the DNA relaxation activity of topoisomerase I. These findings indicate that the interaction of c-Abl with DNA is both sequence-selective and structure-dependent.

The c-Abl protein is a member of the c-Src nonreceptor tyrosine kinase family that differs from c-Src protein by the presence of a large C-terminal part containing several functional domains such as an actin binding site (1), the nuclear localization signal (2, 3), and a DNA-binding domain (4). Mutations within the large C-terminal domain are sufficient to activate the transforming potential of c-Abl (5). The DNA-binding activity is regulated by the cdc2 kinase phosphorylation during the cell cycle (6). During mitosis, phosphorylation of Ser-852 and Ser-883 residues located in the c-Abl DNA-binding domain inhibits the binding of c-Abl to DNA (4). The c-Abl protein can also negatively regulate cell growth via a p53-dependent pathway (7, 8). The c-Abl tyrosine kinase activity is regulated during the cell cycle by interaction with the retinoblastoma protein p105^{RB}. The interaction occurs during G2, M, and G1 phases between the C-pocket of the hypophosphorylated form of p105^{RB} and

the ATP-binding site of c-Abl tyrosine kinase domain and abolishes the tyrosine kinase function (9, 10).

One identified c-Abl tyrosine kinase nuclear substrate is the C-terminal repeat domain (CTD) of the RNA polymerase II large subunit, which is phosphorylated on the tyrosine residues of the 52 hepta-amino acids repeats (11). This tyrosine phosphorylation is proposed to participate in the transition from initiation to elongation phases of the basal transcription complex, as observed for the CTD phosphorylation on serine and threonine residues (12).

Although it is known that c-Abl binds to DNA, the molecular basis of the interaction remains largely unclear. The human c-Abl protein was first proposed to bind to the palindromic EP element of the hepatitis B virus enhancer (13, 14), but this interaction was not confirmed by other authors (15, 16). For example, the activation of c-myc transcription by c-Abl (17) is independent of the presence of the DNA-binding domain of c-Abl or of the EP-like element in the c-myc promoter (15). More recently, Miao and Wang (16) proposed that the murine c-Abl DNA-binding domain interacts with A/T-rich stretches independently of a sequence-specific consensus binding site. They proposed that the c-Abl DNA-binding domain resembles that of the HMG¹ proteins. Therefore, the nature of c-Abl binding sites on DNA remains an area of ongoing controversy, and it is the issue we address here.

Using CASTing experiments, we show that the purified DNA-binding domain of c-Abl recognizes selectively the

[†] This work was supported by the INSERM, the Université de Lille II, the Association pour la Recherche contre le Cancer, and the Ligue Nationale contre le Cancer, comité du Nord. M.-H.D.-C. is supported by the Institut de Recherches sur le Cancer de Lille, the Conseil Régional du Nord-Pas de Calais, and the Association pour la Recherche contre le Cancer. M.H. was supported by the CHRU de Lille, the Conseil Régional du Nord-Pas de Calais, the Institut de Recherches sur le Cancer de Lille, and the Ligue Nationale contre le Cancer.

^{*} Author to whom correspondence should be addressed.

[‡] INSERM.

[§] Present address: IRIBHN, Erasme Hospital, U.L.B., 808 Route de Lennik, B-1070 Brussels, Belgium.

^{||} Centre Oscar Lambret.

consensus binding site A^A/cAACAA^A/c but not necessarily the A/T-rich tracts as reported by Miao and Wang (16). Using a variety of biochemical methods to investigate the DNA-binding properties of c-Abl, we show that the minimal binding domain preferentially binds to distorted DNA structures such as four-way junctions and bubble DNA but, however, does not induce DNA bending. In addition, we present experimental data suggesting that the protein recognizes its consensus sequence via the minor groove of the double helix. On the one hand, c-Abl exhibits HMG-like properties in terms of sequence selectivity and preferential binding to distorted DNA structures, however, on the other hand, c-Abl differs from HMGs because it does not bend DNA. Finally, it is shown that the protein potentiates the topoisomerase I-mediated relaxation of plasmid DNA.

MATERIALS AND METHODS

Oligonucleotides. The oligonucleotide sequences are presented in Figure 4. All oligonucleotides were purchased from Eurogentec and purified. Prior to the annealing reaction, the desired oligonucleotide was 5'-end-labeled using [γ -³²P]ATP in the presence of T4 polynucleotide kinase.

Expression and Purification of the Human c-Abl DNA-Binding Domain in Bacteria. The [NcoI 2373–XmaIII 2912] sequence encoding c-Abl DNA-binding domain was fused at the 3' end to an (His)₇ coding oligonucleotide and cloned in the pAlter-EX1 plasmid (Promega) or the pTRC-99A vector (Pharmacia). The pTRC-99A-NX-HIS construct was used to obtain a bacterially expressed NX-HIS protein. The NcoI–MluI DNA fragment was subcloned into pGEX-2T vector between the NcoI site (which we have previously substituted for the original BamHI site) and the MluI site. The GST-NX-HIS fusion protein was expressed in *E. coli* strain JM109 by induction with IPTG.

The GST-NX-HIS protein was purified using Ni columns in denaturing conditions (InVitrogen) and then refolded to a native structure by gradual dialysis against the EMSA incubation buffer (50 mM HEPES, pH 7.9, 100 mM KCl, 5 mM EDTA, 1 mM PMSF, 20% glycerol). The protein preparation was analyzed by SDS–PAGE. The capacity of the proteins to bind to DNA was verified via an interaction with DNA–cellulose, prior to be used in the CASTing experiments.

Selection of Binding Sites by CASTing. CASTing experiments were performed as described by Harley et al. (18) using 5 μ g of double-stranded R62 oligonucleotide (CAGGTCAGTTCAGCGGATCCTGTCG(N)₁₂GAGGC-GAATTCAGTGCAACTGCAGC) mixed with 50 ng of the purified GST-NX-HIS protein in 100 μ L of binding buffer (20 mM HEPES, pH 7.9; 100 mM KCl, 20% glycerol, 10 mg/mL BSA, 1 mM PMSF) in the presence of aprotinin (1 μ g/mL) and 0.5 μ g of poly(dI-dC)•(dI-dC). The mixture was incubated for 30 min at 4 °C under mild agitation, 25 μ L of pre-equilibrated Ni beads was then added, and the incubation was continued for 15 min at 4 °C. The Ni beads were washed 5–10 times with the binding buffer. After a brief

centrifugation, bound oligonucleotides were isolated from the pellet, which was resuspended in 50 μ L of PCR buffer (Appligene) and incubated at 95 °C for 5 min. The selected oligonucleotides in the supernatant were subjected to 10–20 cycles of PCR (1 min at 94 °C, 1 min at 65 °C, and 2 min at 72 °C) in a final volume of 100 μ L of PCR buffer containing 15 pM of the reverse and forward primers. An aliquot was taken out and analyzed on a 3% agarose gel to verify the efficiency of the PCR. Nine rounds of binding and amplification were undertaken with an increasing stringency using reduced protein concentrations and increasing KCl (up to 300 mM) and poly(dI-dC)•(dI-dC) (up to 5 μ g per reaction). At last, the selected oligonucleotides were cloned using the BamHI and EcoRI restriction sites (underlined in the R62 sequence) in the pBS_{II}SK plasmid and sequenced.

Electromobility Shift Assays. The cloned sequences isolated by CASTing were 3'-end-labeled with [α -³²P]dCTP (3000 Ci/mM, Amersham) using the Klenow fragment of DNA polymerase I, whereas the synthetic oligonucleotides were 5'-end-labeled with T4 polynucleotide kinase and [γ -³²P]ATP (3000 Ci/mM, Amersham). The labeled probes [10 000 cpm (1–5 ng)] were incubated with the GST-NX-HIS protein in EMSA binding buffer [50 mM HEPES, pH 7.9; 100 mM KCl; 1 mM DTT; 5 mM MgCl₂; 0.1 mM EDTA; 20% glycerol; 500–1000-fold poly(dI-dC)•(dI-dC); 1 mM PMSF], in the presence or absence of a cold competitor, prior to being electrophoresed on a 6% polyacrylamide gel in a 0.7× TAE buffer (4.7 mM Tris-HCl; 2.3 mM sodium acetate; 0.7 mM EDTA; pH 7.9) at 100 V for about 2 h at room temperature. Dried gels were scanned with the PhosphorImager.

DNase I Experiments. DNase I footprinting were performed exactly as described (19). The XbaI–XhoI labeled fragment of pBS_{II}SK-05, 3'-labeled on one or the other of the two complementary strands, was incubated with the purified protein and then subjected to the nicking activity of DNase I (Worthington). DNA cleavage products were resolved by electrophoresis under denaturing conditions in polyacrylamide gels (8% acrylamide containing 8 M urea). Electrophoresis was performed for ~2 h at 85 W in TBE buffer (100 mM Tris-HCl; 100 mM boric acid; 2 mM EDTA; pH 8.3). Gels were then dried and analyzed on the PhosphorImager.

Southwestern Blot Analysis. The fusion proteins were analyzed on a 15% SDS–PAGE and transferred onto a nitrocellulose membrane using a graphite electroblotter (Millipore). The membrane was blocked in 5% (w/v) nonfat dry milk containing buffer (50 mM Tris, pH 7.5, 50 mM NaCl; 1 mM EDTA; 1 mM DTT) for 1 h at 42 °C and then washed three times in binding buffer (10 mM Tris, pH 7.5; 50 mM NaCl; 1 mM EDTA; 1 mM DTT) before incubation for 1 h at room temperature in 10 mL of binding buffer with the 05 selected sequence as probe (2 × 10⁶ cpm). Finally, the nitrocellulose membrane was washed three to five times to avoid nonspecific interactions and analyzed with the PhosphorImager.

Melting Temperature Studies. Melting curves were measured using a Uvikon 943 spectrophotometer coupled to a Neslab RTE111 cryostat. Samples were placed in the thermostatically controlled cell holder (10 mm path length), and the quartz cuvettes were heated by circulating water.

¹ Abbreviations: HMG, high-mobility group; 4W, four way; *E. coli*, *Escherichia coli*; CASTing, cyclic amplification of sequence targeting; EMSA, electromobility shift assay; IPTG, isopropyl β -D-(–)-thiogalactopyranoside; GST, glutathion S-transferase; DTT, dithiothreitol; SDS, sodium dodecyl sulfate; bp, base pair(s); nt, nucleotide(s).

The measurements were performed in BPE buffer (6 mM Na_2HPO_4 ; 2 mM NaH_2PO_4 ; 1 mM EDTA; pH 7.1). The temperature inside the cuvette was monitored by a thermocouple in contact with the solution. The absorbance at 260 nm was measured over the range 20–100 °C with a heating rate of 1 °C/min. The melting temperature T_m was taken as the midpoint of the hyperchromic transition.

Circular Permutation Assay (20). The *Bam*HI–*Eco*RI fragment of the pAlter-EX1 polylinker was cloned in tandem at both ends of the 05 selected sequence to give the pBEND-05 vector. A series of 141 bp fragments containing the 05 protein binding sequence at various positions was obtained upon digestion of the plasmid with different restriction enzymes (see Figure 9A). The fragments were 3'-end-labeled, purified, and used in EMSA.

Ring Closure Assays. The protocol used for the circularization experiments has been detailed recently (21). Briefly, 191, 246, and 302 bp fragments were incubated in the presence of the GST-NX-HIS or NX-HIS fusion proteins in 1 × ligase buffer for 20 min at 30 °C prior to the addition of 1 unit of T4 DNA ligase (Life Science) for 15 min at room temperature and then denatured by heating for 15 min at 65 °C. Samples were incubated with 30 units of exonuclease III (Boehringer) for 1 h at 37 °C. Ligation products were resolved by electrophoresis on a 5% polyacrylamide gel.

DNA Relaxation Experiments. Supercoiled pBEND05 DNA (0.77 μg) was incubated with the GST-NX-HIS protein for 10 min in relaxation buffer (50 mM Tris, pH 7.8; 50 mM KCl; 10 mM MgCl_2 ; 1 mM DTT; 1 mM EDTA). Four units of calf thymus DNA topoisomerase I (TopoGen Inc.) was then added to each tube, and the reaction was continued for 1 h at 37 °C prior to the addition of SDS to 0.25% and proteinase K to 250 $\mu\text{g}/\text{mL}$. DNA samples were electrophoresed in a 1% agarose gel. Gels were stained with ethidium bromide (1 $\mu\text{g}/\text{mL}$), washed, and photographed under UV light.

RESULTS

Sequence-Specific Binding Site. To purify the DNA-binding domain of c-Abl protein, we cloned the [*Nco*I 2273–*Xma*III 2912] fragment of c-Abl cDNA in the pGEX-2T vector as described under Materials and Methods. This fragment encodes the c-Abl DNA-binding domain which we have previously delimited from the capacity of deletion mutants expressed in reticulocyte lysates to bind to DNA cellulose. The purified GST-NX-HIS protein binds to DNA cellulose in the same manner as does the large C-terminal part of c-Abl protein containing the DNA-binding domain. The effect is attributable to the NX-HIS part of the bacterial protein since the GST protein is known and has been shown not to bind to DNA. This [*Nco*I 2273–*Xma*III 2912] fragment of the human c-Abl c-DNA spans the minimal DNA-binding domain encoding sequence of the murine c-Abl cDNA. This domain is localized between the kinase domain and the actin-binding site in the C-terminal portion of c-Abl protein (4). The protein GST-NX-HIS, which can be produced in large quantity in bacteria and can be highly purified, has been used for the different experiments described below. The c-Abl DNA-binding domain does not contain any classical DNA-binding motifs but is characterized by its prolin-rich constitution.

Five computer programs were employed to predict the secondary structure of the DNA-binding domain of c-Abl. The agreement between the different methods is satisfactory. They all predicted roughly the same structure: the DNA-binding domain of c-Abl would be composed of three α -helices and three potential β -sheets separated by coiled regions of variable lengths (data not shown).

We used cyclic amplification of sequence targeting (CASTing) experiments to identify the consensus site preferentially recognized by the c-Abl purified DNA-binding domain. For this purpose, the GST-NX-HIS fusion protein was incubated with 62-mer oligonucleotides containing 12 random base pairs. The protein-bound sequences were trapped with Ni beads, amplified by PCR, and used for the following round. After nine rounds of selection–amplification, the ability of the selected oligonucleotides to bind the protein was verified by EMSA and the last amplicones were cloned and sequenced. Forty-one sequences showed a high homology and were used to display the consensus binding site $\text{A}^{\text{A/C}}\text{AACAA}^{\text{A/C}}$ (Figure 1). All of these sequences contain the AAC triplet sequence. Many sequences, including the six selected sequences not used for the determination of the consensus sequence (bottom of Figure 1), contain A·T-rich stretches, but they are shorter than those found by Miao and Wang with murine c-Abl protein (16). The size of the selected sequences varies from 11 to 30 bp and 15 of the 41 selected sequences, as well as the nonselected but A·T-rich clone 25, encompass more than 12 bp. This variation in the length of the selected sequences may be due to Taq-polymerase mismatches during PCR amplification and to the fact that the GST-NX-HIS fusion protein preferentially selected these longer DNA sequences containing also its sequence-specific binding site.

We tested the ability of GST-NX-HIS fusion protein to bind to a few selected sequences (01, 02, 03, 05, 10, and 17) or to the control SK that failed to contain any potential binding site in EMSAs using both radiolabeled *Bam*HI–*Eco*RI DNA probes (32 or 37 bp) and longer *Xba*I–*Xho*I fragments (73 or 78 bp) (Figure 2A). Interestingly, the fusion protein GST-NX-HIS binds very weakly to the *Bam*HI–*Eco*RI probes (data not shown), whereas it binds tightly to the longer *Xba*I–*Xho*I probe. The observations that (i) the fusion protein binds more tightly to 73-mer sequences than to 32-mers containing the same binding core and (ii) the sequences selected by CASTing are frequently longer than the 12 bp randomized sequence suggest that the protein specifically recognizes a central core and that the flanking sequences likely contribute to reinforce the affinity of the protein for its sequence-specific target site. Protein–DNA complexes were observed with probes 01, 02, 03, and 05, to a much lower extent with clone 10 but not with clone 17, and the control probe SK that lacks the target site but, however, contains the same flanking sequences (Figure 2B). The *Xba*I–*Xho*I fragment of clone 17, which does not contain the determined consensus binding site, fails to bind to the protein (Figure 2B, lane 12). Figure 2 as well as competition experiments (not shown) demonstrated that probe 05, which contains three juxtaposed AAC motifs, exhibits the highest affinity for the GST-NX-HIS protein. The amount of protein required to see any binding is relatively large (100 nM–1 μM). This can be attributed to the weak affinity of the protein, possibly because (i) the

```

      G A A A A C A A C A A C      O1
A A A A C A A C A C A C A A C A A      O2
      A A A A C A A C A C A C A A C A A      O2
      A A A A A A C A G A C A G G      O3
      A A A A A A C A C A A C      O4
      A A C A A C A A C A A A      O5
      C A A A A C A A G A A A T A A G A A C A A C G T C      O6
C A A A A C A A G A A A T A A G A A C A A C G T C      O6
      A A A A A C A C G A A A C      O7
      G T A A C A A A C G G G      O8
      G G A A A C A A C A A C      O9
      G A A G C A A A A C A G      10
      A A G C A A A A C A A C G      11
      A C A A C A A C A A C A      12
G G G A A A A C C A A C G A C A A G C A A A A C A A C G      13
      A A G C A A A A C A A C G      14a
      A A A A A C A C G A A A C G A C A A A T C A A A T G      14b
      C A C A A A C A A C A A G A C A C G A      14c
      A A A A C A A G A A A T A A G A A C A A C G T C      15
A A A A C A A G A A A T A A G A A C A A C G T C      15
      C A A C A A A G A A G C A C G A C A A C G A C C      16
      C C A A A A C G A C C A G A A A A A T A G G      18
      C A C A T A A A C A G G      19a
      G T A A G A A C A A G A A A C A A A C A      19b
      G A A C A G C A A A A C      20
      A A A G C A A C A A C G C A      21
      C A A A A A A C G A G A G T      22a
      C A A A C A A C A A C G      22b
C G T G G A A A A C C A A C G A C A A G C C A A A A C A A C      22c
      A A A A C C A A C G A C A A G      23
      G G A A C A A A A A G G      24a
      G C A G C A A C A T G G      24b
      G C T A A G A A C A A G A A A C A A A C A      27a
      G A A A A C A A A C G A A      27b
      G G A A A A C A G C G A      28
      A A A C A T A A G A A A G A A C A A G A A      29
      A G C A C A T A A A C      30
      A A A A C A G A A A A C      32
      G C A A A C A A A C A G G      34
      A A A A C A A C A A T A      35
      G A A A C A A C A A T A C      37
      C A A A A C A A A A T G      38
      G A A C A A C A C A A A      39
      G A C A A A C A A A A G G      40

```

Total Number	28	36	42	44	44	44	44	43	43	42	36	30
--------------	----	----	----	----	----	----	----	----	----	----	----	----

G	8	3	6	5	0	0	0	4	7	9	10	3
A	13	24	31	28	44	44	0	39	31	11	20	21
T	1	0	2	0	0	0	0	0	2	0	0	2
C	7	9	3	11	0	0	44	0	3	22	6	4

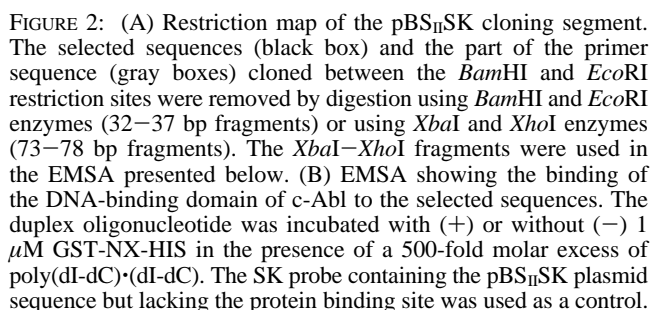
%	46	67	74	63	100	100	100	91	72	52	55	70
Consensus	A	A	A	A	A	A	C	A	A	C	A	A
	G	C		C						A	G	
	29	25		25						26	28	
	C									G	C	
	25									21	17	

```

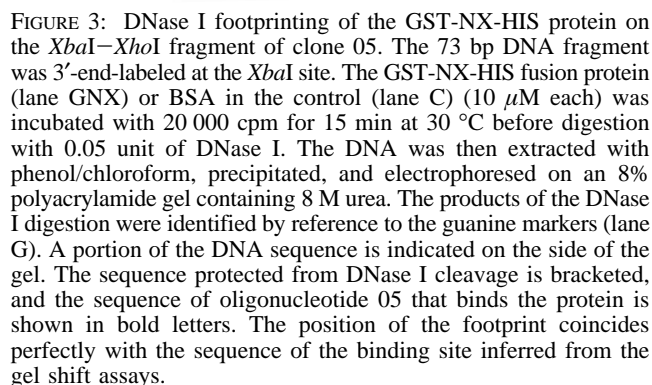
G T G A G G A A A A G G      17
T A A A G T A A A G G T A G G A A A G C A C A      25
C A C A A A T A A A A      26
G C C A A A T C A G G C      31
T G G T A C A C G G G C      33
C C A T A T A C T C T G      36

```

FIGURE 1: Oligonucleotides 01–40 were sequenced after nine rounds of selection–amplification of the 62-mer oligonucleotides containing the random 12 bp. The upper sequences were aligned to show the sequence homology and were used for determining the consensus. The six lower sequences that do not contain an AAC triplet were not used for determining the consensus binding sequence.



The exact position of the protein binding site within probe 05, which presents the highest binding ability for GST-NX-HIS, was located by means of DNase I footprinting. As shown in Figure 3, the sequence protected from cleavage by the nuclease coincides with the sequence identified in the CASTing experiments. There is no binding to the vector sequences or to the primers and cloning site sequences. Therefore, the EMSA and footprinting data concur that the protein binds selectively to 5'-AAC (GTT)-containing sequences.



First, we compared the binding of GST-NX-HIS to linear DNA and to DNA molecules of the same length containing a large single-strand loop (hereafter referred to as bubble DNA) or four-way junction (4W). The four-way junction (also termed Holliday DNA) was proposed to be the central intermediate in homologous recombination events and DNA replication (25). We used two types of 4W junction DNAs,

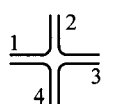
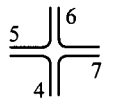
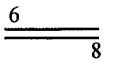
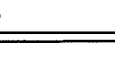


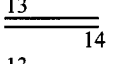
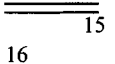
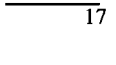
nucleotide sequence	structure	name
1: CTGGTCTACCAGATTACCCCGTGTGAGAGCGCTCGCC 2: GGCAGAGCGCTCTACACGGCCCTCCGCCAGCTGCGTG 3: CACGCAGCTGGGCGGAGGGGCGCGGACGTTAACCGTCC 4: GGACGGTTAACGTCCGCGCGGGTAATCTGGTAGACCAG		4Wa
4: GGACGGTTAACGTCCGCGCGGGTAATCTGGTAGACCAG 5: CTGGTCTACCAGATTACCCCAACAAAAGAGCGCTCGCC 6: GGCAGAGCGCTCTTTTGTGTTTTTGTCCAGCTGCGTG 7: CACGCAGCTGGGCAAAAAGCGCGGACGTTAACCGTCC		4Wb
6: GGCAGAGCGCTCTTTTGTGTTTTTGTCCAGCTGCGTG 8: CACGCAGCTGGGCAAAAACAACAAAAGAGCGCTCGCC		DUP-38mer
9: AGACTAGTGGATCCTGTCTAAAAACAACAACACCGATACCGTCGACCTCGAG 10: CTCGAGGTCGACGGTATCGTGTTTGTTGTTTGTAGACAGGATCCACTAGTCT		DUP-56mer
9: AGACTAGTGGATCCTGTCTAAAAACAACAACACCGATACCGTCGACCTCGAG 11: CTCGAGGTCGACGGTATCGCCCCCCCCCCCCCAGACAGGATCCACTAGTCT		BBL-56mer (AAC)
11: CTCGAGGTCGACGGTATCGCCCCCCCCCCCCCAGACAGGATCCACTAGTCT 12: AGACTAGTGGATCCTGTCTTGGATCCGCCGGCTGCAGCGATACCGTCGACCTCGAG		BBL-56mer (SK)
13: GGCAGAGCGCTCTTCCGCCGCCCGCCAGCTGCGTG 14: CACGCAGCTGGGCGGGGCGGGGCGGGGAGAGCGCTCGCC		IC-38mer
13: GGCAGAGCGCTCTTCCGCCGCCCGCCAGCTGCGTG 15: CACGCAGCTGGGCGGGGCGGGGCGGGGAGAGCGCTCGCC		GC-38mer
16: ACTAGTGGATCCCCGGGCTGCAGGATCGATACAGCTT 17: AAGCTGTATCGATCCTGCAGCCCGGGGATCCACTAGT		SK-38mer

FIGURE 4: Oligonucleotide sequences and structures. The four-way junction DNA 4Wa (used as a control and derived from the 30-mer four-way junction λ (22) and 4Wb (which contains the consensus binding site) were obtained by annealing oligonucleotides 1, 2, 3, and 4 or oligonucleotides 4, 5, 6, and 7, respectively. The 38-mer duplex DNA containing the consensus binding site was obtained by annealing oligonucleotides 6 and 8. The 56-mer duplex containing the consensus binding site was obtained from oligonucleotides 9 and 10. Oligonucleotides 9 and 11 (containing the AAC repeat) and 11 and 12 (containing sequences of the pBS_{II}SK vector polylinker used as control) gave the so-called bubble DNAs (BBL) containing an 18 base single-strand loop on one side and a 15 base single-strand loop on the other side. For the competition assays, the same oligonucleotides were used as well as mutants of the 38-mer duplex which possess I·C (oligonucleotides 13 and 14) or G·C bp (oligonucleotides 13 and 15) in place of the A·T bp in the consensus binding site. Oligonucleotides 16 and 17 were annealed to obtain the SK 38-mer control for the DNA thermal denaturation studies.

one containing the potential c-Abl binding site (named 4Wb) and another lacking the c-Abl binding sites but containing a potential low affinity SRY site (4Wa) (22). The two DNAs possess one identical strand, which was ³²P-labeled prior to annealing to the three other specific oligonucleotides. In EMSA, 4Wb containing the consensus binding site showed a slightly better affinity for GST-NX-HIS (Figure 5A, lanes 8–14) than 4Wa (lanes 1–7). In contrast, many fewer complexes are formed when using the linear duplex (lanes 15–20). The 38-mer duplex, containing the same ³²P-labeled oligonucleotide as 4Wb, apparently binds much more weakly to the protein than the branched oligonucleotides.

Second, we employed 56-mers DNA with a duplex structure or presenting a large 18 base long single-stranded loop. Here again, we found that the protein binds much more tightly to the bubble DNA (BBL) than to the corresponding linear duplex (DUP). No significant differences were found whether or not the single-stranded region of the bubble DNA contains the consensus binding site (data not shown).

Competition experiments with unlabeled linear and bubble DNA molecules were performed to gain further understanding of the selectivity of the protein for linear versus distorted DNA species. As shown in Figure 5B, the addition of the distorted DNA species (BBL or 4W) but not the orthodox double-helical molecules inhibits the formation of the protein–DNA complexes. Unexpectedly, the labeled 56-mer duplex probe bound to GST-NX-HIS is displaced by the unlabeled duplex DNA but not by the unlabeled bubble

DNA or by the 4Wb DNA (Figure 5C). This may be explained by the fact that the c-Abl DNA-binding domain could contain two subdomains with relative independent binding, one interacting with linear DNA in a sequence-dependent manner and the other interacting with distorted DNA in a structure-dependent manner.

Protein analysis confirms that linear and distorted DNA bind to the GST-NX-HIS protein. Indeed, the 46 kDa GST-NX-HIS fusion protein identified on a Coomassie Blue-stained gel (Figure 6A) and on an immunoblot with anti-GST antibody (Figure 6B) is also visible on southwestern blot using the *Xba*I–*Xho*I fragment of clone 05 (not shown) or the 4Wb as probe (Figure 6C).

We also demonstrated that the radiolabeled deletion mutants of c-Abl C-terminal domain, expressed in reticulocyte lysates, interact with bubble-linked agarose beads and with linear 56-mer duplex-linked beads. These two experiments confirm that the binding to either specific sequence and distorted DNA structures is due to the GST-NX-HIS protein and not to a potential copurified protein.

Interaction with Single-Stranded DNA. EMSAs with single-stranded oligonucleotides were undertaken to determine whether the tight binding of the protein to the bubble DNA arises from the recognition of a specific structure or from a preferential interaction with a single-strand region of the DNA. The 56-mer oligonucleotides that have been used to form the linear duplex and the bubble DNA were employed (see Figure 4). The protein does not form

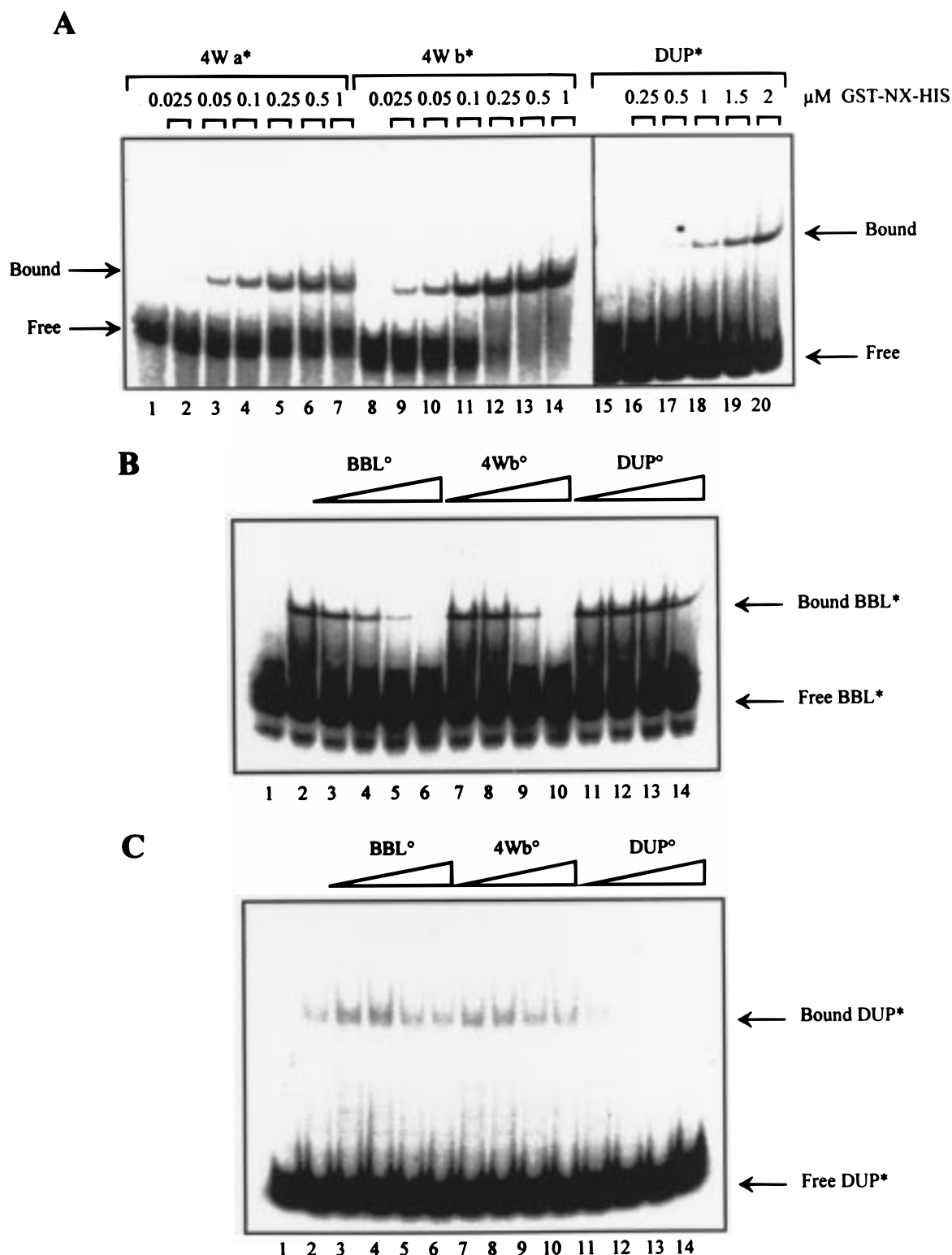


FIGURE 5: (A) EMSA using the two four-way junction DNA 4Wa (lanes 1–7) and 4Wb (lanes 8–14) or the consensus 38 bp duplex DNA (DUP, lanes 15–20). The ^{32}P -labeled DNA was incubated with various amounts of the GST-NX-HIS protein in the presence of a 500-fold molar excess of poly(dI-dC)•(dI-dC). (B) EMSA for the binding of GST-NX-HIS to the linear duplex (DUP), bubble (BBL), and four-way junction (4W) DNAs. GST-NX-HIS (200 nM) was incubated with bubble DNA (BBL*) in the absence (lane 2) or presence of cold bubble DNA (BBL°, lane 3, 10-fold; lane 4, 50-fold; lane 5, 200-fold; lane 6, 1000-fold), cold 4Wb DNA (4Wb°, lane 7, 10-fold; lane 8, 50-fold; lane 9, 200-fold; lane 10, 1000-fold), or the 56-bp duplex DNA (DUP°, lane 11, 10-fold; lane 12, 50-fold; lane 13, 200-fold; lane 14, 1000-fold) in the presence of poly(dI-dC)•(dI-dC). (C) The ^{32}P -labeled duplex DNA (DUP*) probe was incubated with 1 μM GST-NX-HIS fusion protein and the competitor DNA as in (B).

complexes with oligonucleotides 9 and 11, which contain the three AAC triplets and the cytosine repeat, respectively. In contrast, complexes were identified with oligonucleotide 10, which possesses three juxtaposed GTT triplets (Figure

7). However, the extent of binding is much lower than that observed with the bubble DNA. We are inclined to believe that the protein recognizes a distorted structure rather than a single-strand sequence. GTT oligonucleotide 10 may adopt

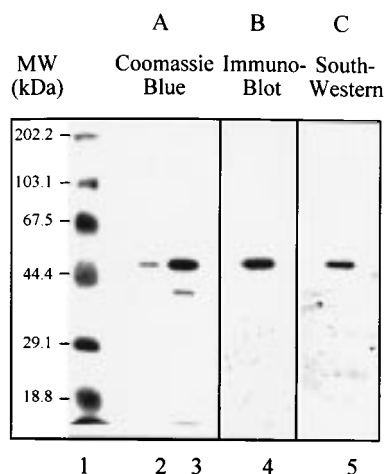


FIGURE 6: Southwestern blot analysis of GST-NX-HIS binding to DNA. (A) The molecular weight markers (lane 1) and 5 or 20 μ g of GST-NX-HIS (lanes 2 and 3) were stained with Coomassie Blue. (B) After electrotransfer on a nitrocellulose membrane, the identity of GST-NX-HIS protein (20 μ g) was verified using a polyclonal anti-GST rabbit antibody (lane 4). (C) For the southwestern blot analysis, the immobilized GST-NX-HIS protein (20 μ g) was revealed for its ability to retain the 05 *XbaI*–*XhoI* fragment (not shown) or the 4Wb probe (lane 5).

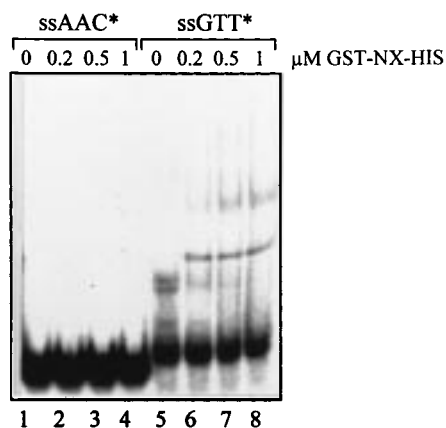


FIGURE 7: Interaction of GST-NX-HIS with single-stranded DNA. EMSA were performed with oligonucleotides 9 and 10 containing the AAC and GTT triplet repeat, respectively. In both cases, the oligonucleotide was incubated in the absence and presence of the fusion protein at the indicated concentration.

a folded conformation that would explain the poor binding to the protein.

DNA Thermal Denaturation Studies. The ability of the protein to alter the thermal denaturation profile of DNA is used as another indication of its propensity to bind selectively to AAC-containing sequences and to distorted DNA structures. We measured the change of the absorbance at 260 nm as a function of the temperature for a series of DNA and polynucleotides of different base compositions and structures in the absence and presence of the GST-NX-HIS protein. In each case, monophasic melting curves were observed. The protein is resistant to heat denaturation in the presence of DNA. The T_m and ΔT_m (T_m of the protein–DNA complex – T_m of the free DNA) values are collated in Table 1. A large increase in the T_m of nucleic acids is observed with the protein using the AAC-containing 38-mer and 56-mer, whereas there is little or no change with nonspecific sequences including poly(dA–dT)•(dA–dT). Similarly, the stabilization of the DNA structure by binding of

Table 1: Variation in Melting Temperature (ΔT_m)^a

	T_m		ΔT_m
	DNA	protein–DNA complex	
poly(dA–dT)•(dA–dT)	41.0	42.0	1.0
poly(dI–dC)•(dI–dC)	36.0	37.0	1.0
SK 38-mer (nonspecific)	57.0	59.0	2.0
AAC-containing 38-mer	55.2	60.5	5.3
AAC-containing 56-mer	57.4	69.5	11.6
four-way DNA 4Wa	39.8	51.8	12.0
four-way DNA 4Wb (AAC)	41.0	55.5	14.5
bubble DNA (SK)	45.7	56.6	10.1
bubble DNA (AAC)	45.0	61.1	16.1

^a The indicated T_m and ΔT_m values (in $^{\circ}$ C) were obtained from first-derivative plots.

the protein is significantly higher with the AAC-containing four way (4Wb) and bubble DNAs compared to the corresponding structures lacking the AAC motif. The results are in agreement with the gel retardation assays and confirm that the protein does exhibit a considerable preference for AAC-containing sequences and unusual DNA structures.

Minor Groove Recognition. The recognition of the consensus binding sequence by c-Abl can occur via the major or the minor groove of the double helix. Most proteins use the major groove for interaction with specific sequences in DNA; however, there exist numerous proteins, such as the purine repressor (PurR), TBP, and proteins of the HMG family (LEF-1, SRY), that interact essentially within the minor groove of DNA (26). The potential analogy between c-Abl and HMG proteins in terms of DNA recognition prompted us to examine this question. First, we investigated the effect of the AT-specific minor groove binding antibiotic netropsin on the complex formation between GST-NX-HIS and the 05 *XbaI*–*XhoI* sequence. EMSA and DNase I footprinting experiments have revealed that the fusion protein binds very tightly to this sequence. The antibiotic should affect the recognition process if the protein–DNA interaction occurs within the minor groove but not if it takes place in the opposite major groove. This approach has previously been employed with success to characterize minor versus major groove binding proteins (27). As shown in Figure 8A, netropsin competes with the formation of the GST-NX-HIS binding to probe 05. A concentration of 1 μ M netropsin is sufficient to reduce the amount of protein–DNA complexes by \sim 50%, and the binding of the protein is almost completely inhibited with 5 μ M netropsin. This suggests that GST-NX-HIS interacts within the minor groove of the double helix.

Second, we resorted to a strategy previously used to evidence minor groove binding of HMG proteins to DNA (28) and which consists of replacing the A•T bp present in the consensus binding site with I•C or G•C bp. If the interaction between c-Abl and its target sequence takes place in the minor groove, the substitution of A•T bp by I•C bp should not affect the interaction since these two types of base pairs are equivalent in terms of hydrogen bond donors or acceptors in the minor groove. In contrast, the replacement of A•T with G•C bp should hinder the access of the protein since, in this case, the 2-amino group of guanine obstructs the minor groove. Figure 8B shows the respective ability of the three unlabeled 38-mer duplexes containing

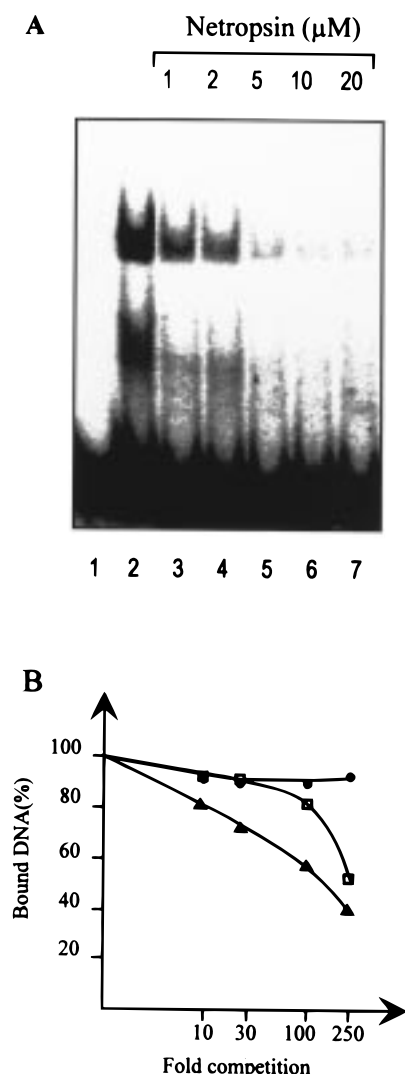


FIGURE 8: (A) EMSA of GST-NX-HIS in the presence of netropsin. The 05 *XbaI*–*XhoI* probe (lane 1) was incubated with 1 μM GST-NX-HIS fusion protein (lane 2) before the addition of increasing concentrations of netropsin (lanes 3–7). (B) Binding plots showing the dissociation of the GST-NX-HIS bound to the 05 *XbaI*–*XhoI* probe in the presence of increasing amounts of the duplex 38-mer oligonucleotide containing A·T (\blacktriangle), I·C (\square), or G·C (\circ) bp in the binding site. The fusion protein (1 μM) was bound to the ^{32}P -labeled 05 *XbaI*–*XhoI* probe before the competitor oligonucleotide was added. The fraction of bound DNA (%) is plotted as a function of the oligonucleotide/[^{32}P]DNA molar ratio. Data are compiled from quantitative analysis of two independent experiments. The x -axis is logarithmic.

A·T, I·C, or G·C bp to compete with the formation of the complexes between GST-NX-HIS and the 05 *XbaI*–*XhoI* probe. The duplex with 10 G·C bp instead of the canonical A·T bp completely fails to inhibit the protein–DNA interaction, whereas the corresponding duplex with 10 I·C bp does interfere with the formation of the complexes. The A·T \rightarrow I·C substitution slightly reduces the binding capacity of the protein but proves to be much more easily tolerated than the corresponding A·T \rightarrow G·C substitution. The I·C data suggest that minor groove interactions are important (the competition with netropsin also supports this view) but are not the sole determinant of binding.

c-Abl Does Not Induce DNA Bending. One of the most typical characteristics of HMG proteins is their ability to induce bending of DNA. For example, SRY and LEF-1 bend

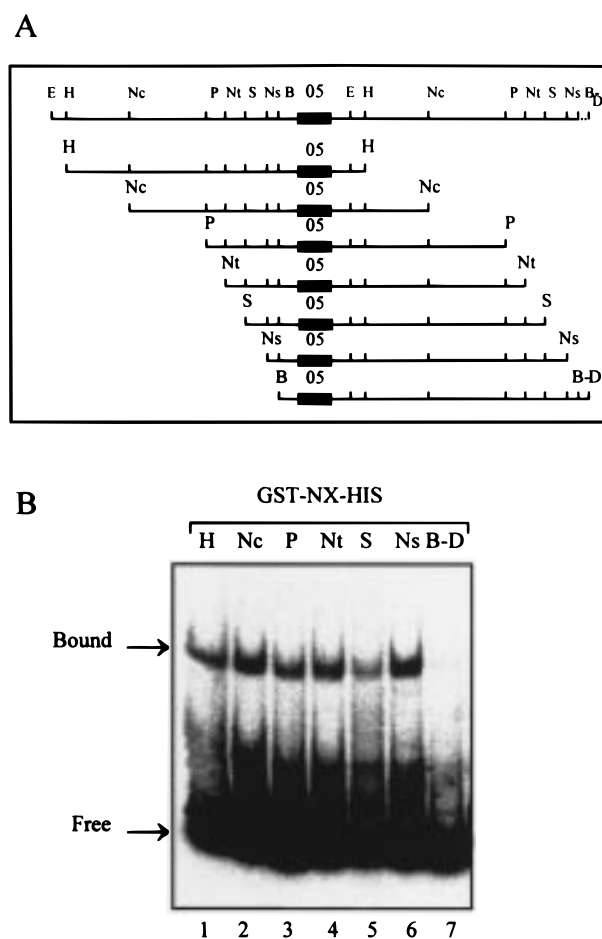


FIGURE 9: Circular permutation analysis of the binding of the GST-NX-HIS fusion protein (1 μM) to DNA fragments containing the 05 binding site. A homologous series of 141 bp restriction fragments (A) was produced by digesting the plasmid pBEND-05 with *HpaI* (H), *NcoI* (Nc), *PstI* (P), *NotI* (Nt), *SphI* (S), *NsiI* (Ns), *BamHI* (B), and *DraII* (D) restriction enzymes. The electrophoretic mobility (B) of various protein–DNA complexes remains unchanged; however, the 05 binding sequence (solid box) is located near the 3' or 5' end of the fragment. In each case, the protein (1 μM) was incubated with DNA probe in the presence of a 500-fold molar excess of poly(dI-dC)·(dI-dC).

DNA by 85° and 130°, respectively (28). It was thus interesting to determine if *c-Abl* can also bend its target sequence, as do HMG proteins. For that, the well-established circular permutation assay was employed (20). The pBEND-05 plasmid was constructed by inserting a polyrestriction site sequence on both the 5' and 3' sides of the 05 *EcoRI*–*BamHI* fragment known to bind tightly to the GST-NX-HIS protein (see Materials and Methods for the plasmid construct). A series of 141 bp fragments was thus generated upon digestion of the plasmid with different restriction enzymes (Figure 9A), and each fragment, 5'- ^{32}P -labeled and purified, was used for the EMSA with the fusion protein. As shown in Figure 9B, the different protein–DNA complexes migrate at the same position in the gel whatever the position of the protein binding site in the sequence. Identical results were obtained with the NX-HIS protein. The very weak binding observed with the doubly digested *BamHI*–*DraII* sequence, where the protein binding site is close to the 3' end, is in accordance with the EMSA data. These experiments strongly suggest that, unlike most HMG pro-

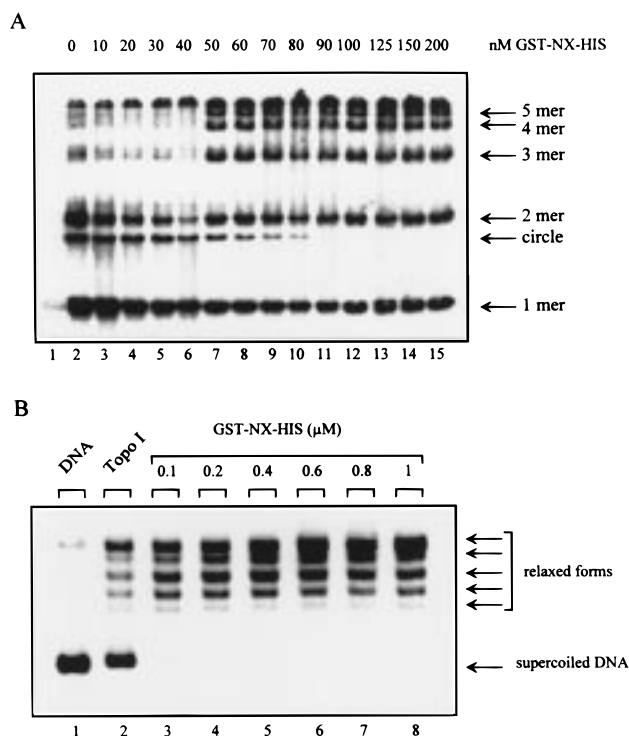


FIGURE 10: (A) Circularization assays. The 5'-end-labeled DNA fragment was incubated with the fusion protein at the indicated concentration in the presence of T4 DNA ligase. The DNA substrate (lane 1) was incubated with 1 unit of T4 DNA ligase in the absence (lane 2) or presence of the protein. The band corresponding to circular DNA was identified by its resistance to exonuclease III. (B) Topoisomerase I-mediated relaxation of the supercoiled pBEND-05 plasmid (0.02 μ M) in the absence (lane Topo I) or presence of increasing concentrations of GST-NX-HIS (lanes 3–8). The protein/DNA molar ratios were 5 (lane 3), 10 (lane 4), 20 (lane 5), 30 (lane 6), 40 (lane 7), and 50 (lane 8).

teins, the c-Abl DNA-binding domain does not bend its target sequence.

Next, we examined the effect of the fusion protein on the cyclization of DNA in the presence of T4 DNA ligase. This ring closure assay was first performed with a 121 bp fragment containing the 05 binding sequence in its central region. Such a fragment is too short to circularize by itself, but circles may be formed in the presence of a DNA-bending protein. DNA minicircles as small as 99 bp have been formed in the presence of HMG-1 (29). With both GST-NX-HIS and NX-HIS proteins, multimers were obtained but no exonuclease III-resistant products (i.e., circular DNA) were detected (data not shown). Then, we resorted to longer fragments that are intrinsically capable of forming circular DNA species, and we tested the propensity of the protein to increase or to decrease the percentage of circles formed. Each restriction fragment was obtained from the plasmid pBS_nSK-05 and contained the 05 protein-binding sequence near the center of the fragment. The gel in Figure 10A shows unambiguously that the protein inhibits the formation of the circular DNA species. As the protein concentration increases from 10 to 200 nM, the band corresponding to the circular DNA (as judged from the resistance to an exonuclease III treatment) decreases rapidly. Similar results were obtained when using 246-mer and 302-mer fragments. In each case, the protein inhibits the formation of circular DNA. Therefore, the data from the circular permutation and the ring closure

assays leave no room for doubt that the protein does not induce a bending of DNA.

Effect on Topoisomerase I-Mediated Relaxation of DNA. We have investigated the binding of the fusion protein to the supercoiled pBEND-05 DNA containing the preferred binding sequence. Even with a high protein concentration (2 μ M), the electrophoretic mobility of supercoiled DNA in agarose gels remains unaffected. No topoisomers or protein–DNA complexes could be seen (data not shown). We then examined the effect of the protein on the relaxation of DNA by human topoisomerase I (Figure 10B). In the absence of the fusion protein, the enzyme removes negative supercoils in native DNA (lane 1), and a population of relaxed DNA molecules, characterized by their reduced electrophoretic mobility, is formed (lane 2). In the presence of a concentration of GST-NX-HIS as low as 0.1 μ M (lane 3), the DNA is almost completely converted into relaxed forms. The effect is even more pronounced at higher protein concentrations (lanes 4–8). GST-NX-HIS stimulates the relaxation activity of topoisomerase I. The effect differs markedly from that recently reported with HMG proteins such as HMG-D and HMG-I/Y, which protect from relaxation by topoisomerase I conferred on negative supercoiled DNA (30, 31). Unlike HMG proteins, which constrain negative supercoils, GST-NX-HIS favors the relaxation of DNA. This observation is consistent with the above results, indicating that the protein inhibits the formation of circular DNA molecules. The protein's effect on the ability to topoisomerase activity would be consistent with preferential binding of the protein to relaxed DNA, shifting the equilibrium toward more relaxed than supercoiled DNA.

DISCUSSION

In this paper, we have demonstrated that the DNA-binding domain of c-Abl binds to a specific consensus sequence in DNA and interacts with distorted DNA structures such as synthetic four-way junctions and bubble DNA containing a large single-strand loop. As such, c-Abl behaves as a HMG protein. The sequence selectively targeted by c-Abl is reminiscent of that recognized by HMG-like proteins such as LEF-1 and SRY (5'-AACAAAG and 5'-^A/TAAACA^A/T, respectively). These two HMG-like proteins and many others (e.g., HMG-I and HMG-D) exhibit a selectivity for deformed DNA structures (22, 28, 32). Moreover, the interaction of the c-Abl DNA-binding domain seems to occur via the minor groove of the consensus DNA sequence as observed for HMG family proteins.

Despite these similarities, it is obvious that c-Abl differs from HMG proteins from at least two independent points. First, there is no strict amino acid sequence homology between c-Abl and HMGs. Second, unlike HMG proteins, c-Abl does not induce DNA bending of its specific target sequences. Furthermore, the ring closure assays (Figure 10A) as well as the experiments based on topoisomerase I-mediated relaxation of DNA (Figure 10B) suggest that c-Abl participates with topoisomerase I to relax DNA. This contrasts with the effects of HMG proteins, which participate with topoisomerase I to supercoil DNA (31). Therefore, c-Abl apparently belongs to a distinct class of DNA-binding proteins.

In their paper, Miao and Wang (16) proposed that three HMG-related domains constitute the mouse c-Abl DNA-

binding domain; however, these domains lack the PRGRP amino acid motif crucial for the DNA-binding properties of HMG-I (33) and are devoid of tryptophan, tyrosine, and threonine residues that are almost always found in the DNA-binding domain of HMG and the related HMG-like proteins (34). In other words, given the poor amino acid sequence homology between c-Abl and HMG proteins, we prefer not to consider human c-Abl as a member of the HMG family. The experimental data also support this view.

The DNA-binding domain of human p145^{c-Abl}, which we have located from a series of deletion mutants, appears to be homologous to that previously identified with the mouse p150^{c-Abl} protein (4). The analysis of the sequence specificity of c-Abl shows unambiguously that the human protein exhibits a high preference for sequences containing an AAC motif. Here again, the results are not totally consistent with those reported for the mouse protein. Miao and Wang indicated that p150^{c-Abl} interacts preferentially with AT-rich tracts as do HMG proteins (32) but does not recognize any particular specific sequences. However, it is possible that human and murine c-Abl DNA-binding domains interact with different DNA sequences. In our hands, the purified human DNA-binding domain binds to certain AT-rich sequences but largely prefers AAC-containing sequences. All of the binding data concur that AAC-based sequences are much more preferred than purely AT-rich sequences. In this respect, it is worth noting that the human c-Abl protein was previously shown to bind to the EP palindromic element of the HBV enhancer that corresponds to the sequence CGT-TGCtcgGCAACG (13), which interestingly does not contain AT-rich stretches but two AAC motifs.

The highly conserved AAC motif is important for the protein-DNA recognition process to occur, but it is surely not sufficient. Inclusion of an AAC triplet into an AT-rich sequence cannot guarantee tight binding of c-Abl to that sequence. The nature of the flanking sequences is important as well. We are inclined to believe that, as with many DNA-binding proteins, there are two levels of recognition of DNA by c-Abl: direct recognition of the consensus binding site coupled with an indirect recognition [the so-called digital and analogue readouts (26)]. The same binding ability was observed with the c-Myb minimal DNA binding domain which needs a 3' extension of more than 10 bp to allow binding to the MRE (myb responsive element). This extension of one turn of helix is important for stabilizing the protein-DNA complex and increasing its half-life (35). In some way, the binding properties of c-Abl are reminiscent of those of restriction enzymes such as *FokI*, which recognizes a specific sequence and which cleaves DNA at sites 9 and 13 bp away from the recognition site (36).

On the basis of the similarities between c-Abl and HMG proteins in terms of binding to specific sequences and deformed structures, it is tempting to classify c-Abl as an atypical HMG-like protein, all the more than HMG and c-Abl share the additional property of being regulated by cdc2 kinase phosphorylation during the cell cycle (37). The similitude is even more striking when one considers that, as shown here, the interaction between c-Abl and its target sequence essentially takes place in the minor groove of DNA, just as with HMGs. There is no doubt that c-Abl exhibits the rare property of interacting in the minor groove of DNA as shown for PurR, SRY, and TBP (26). However, the lack

of detectable induced bending of DNA (as judged by both the circular permutation and ring closure assays) is specific to the c-Abl tyrosine kinase. It must be acknowledged that the DNA-binding properties evidenced in this study concern only the fusion proteins GST-NX-HIS and NX-HIS produced in bacteria. Although it is likely that the same properties applied for the full-length c-Abl, the possibility remains that c-Abl, but not the two truncated mutants used in this study, bends DNA. The *NcoI*-*XmaIII* construct obviously encodes the amino acid sequence required for binding to DNA but perhaps lacks the DNA-bending domain. With the drosophila HMG-D protein, it is known that the acidic tail and the basic region are not required for the interaction with DNA, but their presence is necessary to increase the efficiency of constraining preformed negative supercoils and for optimal DNA bending, respectively (31). By analogy, a specific basic domain may exist in the full-length c-Abl tyrosine kinase.

c-Abl has been implicated in DNA recombination and repair (38) and very recently in the regulation of DNA damage-induced apoptosis (39). So far, it is not possible to establish a direct connection between these biological effects and the particular DNA-binding capacity of the protein. However, it would be a strange coincidence if the interaction of c-Abl with DNA was to play no part in the effect of the protein at the recombination level. It is likely that the capacity of c-Abl to recognize a specific AAC-containing sequence as well as a deformed structure is responsible, at least in part, for some of the biological activities of the protein. As shown here, the protein showed a marked preference for distorted DNA conformations, that is, the types of structures frequently encountered in recombination hot spots. Proteins that exhibit a selectivity for the deformed DNA structures (e.g., DNA junctions) have been isolated from a wide variety of sources from eubacteria to mammals, and many of them are involved in site-specific recombination (40). In this context, it is interesting to mention the Tn916 and lambda integrases which both bind to a consensus sequence containing an AAC unit (41). By interacting with such sequences and unusual structures, c-Abl may cooperate with enzymes that mediate recombination. In addition, the interaction of c-Abl with the loop of the transcriptionally active DNA may permit the correct positioning of c-Abl tyrosine kinase within the transcription complex so as to facilitate the phosphorylation of the C-terminal repeat domain of RNA polymerase II.

In conclusion, the present study indicates that c-Abl cannot be considered as a protein of the HMG family and that this protein binds both to specific (AAC)-containing linear sequence and to distorted DNA structures. These unusual properties likely play an important role in the biological activities of c-Abl tyrosine kinase.

ACKNOWLEDGMENT

We thank Dr. E. Canaani for providing us the pSP65BCR-ABL plasmid. We are grateful to Hubert Cordonnier for his assistance with the computer analysis of the sequences isolated from the CASTing experiment and to M.-C. Du-vieuxbourg for the photographs.

REFERENCES

1. McWhirter, J. R., and Wang, J. Y. J. (1993) *EMBO J.* **12**, 1533–1546.
2. van Etten, R., Jackson, P., and Baltimore, D. (1989) *Cell* **58**, 669–678.
3. Wen, S. T., Jackson, P. K., and van Etten, R. A. (1996) *EMBO J.* **15**, 1583–1595.
4. Kipreos, E. J., and Wang, J. Y. J. (1992) *Science* **256**, 382–385.
5. Goga, A., McLaughlin, J., Pendergast, A. M., Parmar, K., Muller, A., Rosenberg, N., and Witte, O. N. (1993) *Mol. Cell. Biol.* **13**, 4967–4975.
6. Kipreos, E. J., and Wang, J. Y. J. (1990) *Science* **248**, 217–220.
7. Goga, A., Liu, X., Hambuch, T. M., Senechal, K., Major, E., Berk, A. J., Witte, O. N., and Sawyers, C. L. (1995) *Oncogene* **11**, 791–799.
8. Sawyers, C. L., McLaughlin, J., Goga, A., Havlik, M., and Witte, O. N. (1994) *Cell* **77**, 121–131.
9. Wang, J. Y. J. (1993) *Curr. Opin. Genet. Devel.* **3**, 35–43.
10. Welch, P. J., and Wang, J. Y. J. (1995) *Mol. Cell. Biol.* **15**, 5542–5551.
11. Baskaran, R., Dahmus, M. E., and Wang, J. Y. J. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 11167–11171.
12. Young, R. A. (1991) *Annu. Rev. Biochem.* **60**, 689–715.
13. Dikstein, R., Heffetz, D., Ben-Neriah, Y., and Shaul, Y. (1992) *Cell* **69**, 751–757.
14. Dikstein, R., Amagi, R., Heffetz, D., and Shaul, Y. (1996) *Proc. Natl. Acad. Sci. U.S.A.* **93**, 2387–2391.
15. Arcinas, M., Sizer, K. C., and Boxer, L. M. (1994) *J. Biol. Chem.* **269**, 21919–21924.
16. Miao, Y.-J., and Wang, J. Y. J. (1996) *J. Biol. Chem.* **271**, 22823–22830.
17. Wong, K. K., Hardin, J. D., Boast, S., Cooper, C. L., Merrel, K. T., Doyle, T. G., Goff, S. P., and Calame, K. L. (1995) *Oncogene* **70**, 705–711.
18. Harley, V. R., Lovell-Badge, R., and Goodfellow, P. N. (1994) *Nucleic Acids Res.* **22**, 1500–1501.
19. Graves, B. J., Jonhson, P. J., and McKnight, S. L. (1986) *Cell* **44**, 565–576.
20. Wu, H. M., and Crothers, D. M. (1984) *Nature* **308**, 509–513.
21. Bailly, C., Minnock, A., and Waring, M. J. (1996) *FEBS Lett.* **396**, 253–256.
22. Ferrari, S., Harley, V. R., Pontiggia, A., Goodfellow, P. N., Lovell-Badge, R., and Bianchi, M. E. (1992) *EMBO J.* **11**, 4497–4502.
23. Giese, K., Amsterdam, A., and Grosschedl, R. (1991) *Genes Dev.* **5**, 2567–2578.
24. Love, J. J., Li, X., Case, D. A., Giese, K., Grosschedl, R., and Wright, P. E. (1995) *Nature* **376**, 791–795.
25. Lilley, D. M. J., and Clegg, R. M. (1993) *Annu. Rev. Biophys. Biomol. Struct.* **22**, 299–328.
26. Travers, A. A. (1995) *Nat. Struct. Biol.* **2**, 615–618.
27. Suzuki, M. (1989) *EMBO J.* **8**, 797–803.
28. Giese, K., Cox, J., and Grosschedl, R. (1992) *Cell* **69**, 185–195.
29. Pil, P. M., Chow, C. S., and Lippard, S. J. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 9465–9469.
30. Nissen, M. S., and Reeves, R. (1995) *J. Biol. Chem.* **270** (9), 4355–4360.
31. Payet, D., and Travers, A. A. (1997) *J. Mol. Biol.* **266**, 66–75.
32. Bianchi, M. E., Beltrame, M., and Paonessa, P. (1989) *Science* **243**, 1056–1059.
33. Geierstanger, B. H., Volkman, B. F., Kremer, W., and Wemmer, D. E. (1994) *Biochemistry* **33**, 5347–5355.
34. Landsman, D., and Bustin, M. (1993) *BioEssays* **15**, 539–546.
35. Gabrielsen, O. S., Sentenac, A., and Fromageot, P. (1991) *Science* **253**, 1140–1143.
36. Yonezawa, A., and Sugaira, Y. (1994) *Biochim. Biophys. Acta* **1219**, 369–379.
37. Reeves, R., Langan, T. A., and Nissen, M. S. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**, 1671–1675.
38. Chen, Y.-Y., Wang, L. C., Huang, M. S., and Rosenberg, N. (1994) *Genes Dev.* **8**, 688–697.
39. Yuan, Z.-M., Huang, Y., Ishiko, T., Kharbanda, S., Weichselbaum, R., and Kufe, D. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 1437–1440.
40. Lilley, D. M. J. (1994) *J. Mol. Recognit.* **7**, 71–78.
41. Lu, F., and Churchward, G. (1994) *EMBO J.* **13**, 1541–1548.

BI973030W